Molecular Evolution of α-Latrotoxin, the Exceptionally Potent Vertebrate Neurotoxin in Black Widow Spider Venom

Jessica E. Garb*1 and Cheryl Y. Hayashi2
1Department of Biological Sciences, University of Massachusetts, Lowell
2Biology Department, University of California, Riverside
*Corresponding author: E-mail: Jessica_Garb@uml.edu.
Associate editor: Todd Oakley

Abstract

Black widow spiders (members of the genus Latrodectus) are widely feared because of their potent neurotoxic venom. α-Latrotoxin is the vertebrate-specific toxin responsible for the dramatic effects of black widow envenomation. The evolution of this toxin is enigmatic because only two α-latrotoxin sequences are known. In this study, ~4 kb α-latrotoxin sequences and their homologs were characterized from a diversity of Latrodectus species, and representatives of Steatoda and Parasteatoda, establishing the wide distribution of latrotoxins across the mega-diverse spider family Theridiidae. Across black widow species, α-latrotoxin shows ≥94% nucleotide identity and variability consistent with purifying selection. Multiple codon and branch-specific estimates of the nonsynonymous/synonymous substitution rate ratio also suggest a long history of purifying selection has acted on α-latrotoxin across Latrodectus and Steatoda. However, α-latrotoxin is highly divergent in amino acid sequence between these genera, with 68.7% of protein differences involving non-conservative substitutions, evidence for positive selection on its physicochemical properties and particular codons, and an elevated rate of nonsynonymous substitutions along α-latrotoxin’s Latrodectus branch. Such variation likely explains the efficacy of red-back spider, L. hasselti, antivenom in treating bites from other Latrodectus species, and the weaker neurotoxic symptoms associated with Steatoda and Parasteatoda bites. Long-term purifying selection on α-latrotoxin indicates its functional importance in black widow venom, even though venterbrates are a small fraction of their diet. The greater differences between Latrodectus and Steatoda α-latrotoxin, and their relationships to invertebrate-specific latrotoxins, suggest a shift in α-latrotoxin toward increased vertebrate toxicity coincident with the evolution of widow spiders.

Key words: venom, α-latrotoxin, Latrodectus, toxin evolution, black widow spider.

Introduction

Venoms are chemically complex secretions produced by some animals for defense or prey acquisition. Several organismal lineages have independently become venomous and exhibit morphological convergence in venom-injecting organs, as well as similarities at the molecular level in their toxin repertoire due to independent recruitment of related genes for venom production (Fry et al. 2009). Venom toxins have drawn enormous scientific attention because of their applications as pharmaceuticals and as probes for isolating cellular receptors (Adams and Olivera 1994; Lewis and Garcia 2003; Veiseh et al. 2007; Williams et al. 2008). Genes encoding venom toxins are also of significant evolutionary interest because of their direct role in organismal fitness and ecological adaptation. Notably, venom molecular evolution is highly dynamic and appears to be shaped by frequent gene duplications and strong diversifying selection, as well as co-evolution and convergence (Duda and Palumbi 1999; Li et al. 2005; Fry et al. 2006; Aminetzach et al. 2009; Binford et al. 2009; Doley et al. 2009).

The order Araneae (spiders) is the largest clade of venomous organisms, but biochemical characterization of their venoms has been restricted to a small sampling of species (Fletcher et al. 1997; Adams 2004; Ushkaryov et al. 2004; Wullschleger et al. 2004), and evolutionary analyses of spider venoms have been limited (Sollod et al. 2005; Binford et al. 2009). A striking example of this dearth of studies is the molecular composition of black widow spider venom, which has been extensively characterized from only one of the 31 Latrodectus species in the family Theridiidae (Latrodectus tredecimguttatus; Knipper et al. 1986; Kiyatkin et al. 1990, Kiyatkin, Dulubova, and Grishin 1993; Dulubova et al. 1996; Volynski, Volkova, et al. 1999). “Black widow” is a common name referring to several Latrodectus species (e.g., L. tredecimguttatus, L. mactans, L. hesperus, and L. variolus), which are widely recognized and feared because of the extreme neurotoxicity of their venom and their abundance in human-inhabited areas (Clark et al. 1992; Muller 1992; Vetter and Isbister 2008). Other Latrodectus species, including the sexually cannibalistic Australian red-back spider (L. hasselti) and highly invasive brown widow (L. geometricus), also inflict painful bites with similar physiological effects to those from black widows but with varying degrees of potency (Muller et al. 1989, 1992; Muller 1992; Isbister and Gray 2003b).

Complementary DNA (cDNAs) for two families of venom protein components—latrotoxins and latrodectins—have been cloned from the venom glands of the Eurasian black...
Latrodectins include only two paralogous sequences, with unclear functional roles (Kiyatkin, Dulubova, Chekhovskaya, et al. 1993; Pescatori et al. 1995). By contrast, numerous studies have shown that latrotoxins are the primary neurotoxins in black widow spider venom (Knipper et al. 1986; Kiyatkin et al. 1990, Kiyatkin, Dulubova, and Grishin 1993; Dulubova et al. 1996; Volynski, Volkova, et al. 1999). The four sequenced latrotoxin paralogs encode long polypeptides (1,200–1,400 amino acids) that share 30–60% amino acid sequence identity and an overall similarity in domain organization (Ushkaryov et al. 2004). Although their general role in stimulating neurotransmitter secretion is also similar, experimental evidence suggests the latrotoxins vary in target specificity. For example, α-latroinsectotoxin and δ-latroinsectotoxin appear to selectively affect insect neurons, whereas α-latruncrustotoxin stimulates secretion from crustacean neurons, but not from certain insects (Fritz et al. 1980; Krasnopernov et al. 1991; Magazanik et al. 1992; Kiyatkin et al. 1995; Elrick and Charlton 1999). The fourth paralog, α-latrotoxin, is a vertebrate-specific neurotoxin that causes the extreme pain following black widow bites (Knipper et al. 1986; Kiyatkin et al. 1990; Volynski, Nosyreva, et al. 1999).

Because of its role in human toxicity and its importance for understanding vertebrate neurosecretion, considerable efforts have focused on the structure–function relationship of α-latrotoxin. The three-dimensional form of α-latrotoxin is a homotetramer resembling a propeller with four “blades” and a central pore, through which calcium ions pass (Orlova et al. 2000). α-Latrotoxin binds to the vertebrate pre-synaptic neuronal receptor neurexin and inserts into the neuronal membrane, becoming a transmembrane channel (Orlova et al. 2000). This leads to an influx of extracellular calcium through α-latrotoxin’s central pore, triggering a massive uncontrolled exocytosis of neurotransmitters. Neurotransmitter release may also be stimulated by α-latrotoxin dimers via intracellular calcium release upon binding to the latrophilin/CIRL receptor (Volynski et al. 2003). The α-latrotoxin monomers have three tandem domains: the wing, body, and head region (320, 694, and 163 amino acids, respectively; Orlova et al. 2000; fig. 1A and B). The wing region is hypothesized to be a receptor-binding domain, whereas the head and body make up the transmembrane channel (Orlova et al. 2000; Ushkaryov et al. 2004).

In contrast to this detailed functional knowledge of α-latrotoxin from L. tredecimguttatus, the evolution of latrotoxins is poorly understood. There is no clear evolutionary link between latrotoxins and any other known protein, and latrotoxins are unknown from spider venoms outside of Latrodectus, suggesting that black widow venom is the product of dramatic molecular evolution. Species in the theridiid genera Steatoda and Parasteatoda can inflict bites with similar, but far less severe, neurotoxic symptoms to those of black widows (Muller et al. 1992; Graudins et al. 2002; Isbister and Gray 2003a). Moreover, median lethal dose values (LD₅₀) in mice vary substantially among venoms from different Latrodectus and Steatoda species (Muller et al. 1989, 1992). The molecular basis for this interspecific diversity may be explained by variation in the phylogenetic distribution, expression, or sequence characteristics of α-latrotoxin. The adaptive significance of a vertebrate-specific toxin in black widow venom is unclear, given that their diet is primarily invertebrate-based, although capture and consumption of small vertebrates by widow spiders is well documented (e.g., geckos, small lizards, snakes, and mice; McCormick and Polis 1982; Hodar and Sanchez-Pinero 2002).

The recent cloning of α-latrotoxin from L. hasselti confirms its presence in other Latrodectus species (Graudins et al. 2012); however, the broader genetic variability and evolution of α-latrotoxin remains largely unknown. Through a combination of genomic polymerase chain reaction (PCR), reverse transcriptase-PCR (RT-PCR) of venom gland cDNA and inverse PCR, we have obtained the coding sequence of the α-latrotoxin gene in its approximate entirety (~4 kb) from divergent representatives of Latrodectus and Steatoda species. We estimated evolutionary relationships of these sequences to other latrotoxin gene family members and have investigated patterns of variability and selection across α-latrotoxin’s structural domains using multiple methods. Further, we have sequenced a portion of α-latrotoxin spanning parts of the wing and body domain from a denser sampling of Latrodectus species. These data were compared with the mitochondrial gene cytochrome c oxidase I (mt COI) to evaluate the relative rate of α-latrotoxin evolution. Our results indicate a strong functional role for α-latrotoxin in a larger set of species than previously recognized, which has implications for the clinical treatment of widow spider bites, as well as for understanding the evolutionary ecology of black widows.

**Results**

**α-Latrotoxin Sequence Variability**

We obtained eight approximately 4 kb α-latrotoxin sequences from divergent Latrodectus species and Steatoda grossa. These sequences exhibited minor length variation, with L. geometricus and S. grossa α-latrotoxin having an overlapping 6 bp insertion relative to all other sequences, one additional 3 bp deletion in L. geometricus, and two separate 3 bp deletions in S. grossa. As previously indicated by southern blots (Danilevich and Grishin 2000) and genomic sequence of α-latrotoxin from L. hasselti (Graudins et al. 2012), all α-latrotoxin gene sequences appear to be intronless. Translations of obtained α-latrotoxin sequences did not contain any unexpected stop codons. We also obtained a 618 bp fragment of α-latrotoxin, spanning part of its wing and body domains, from 41 specimens (sampling 18 Latrodectus and two Steatoda species). Translations of the 618 bp fragment of α-latrotoxin from all Latrodectus species also exhibited no length variation or stop codons, but Steatoda capensis had a 9 bp deletion, and two α-latrotoxin paralogs sequenced from Parasteatoda were either 27 or 15 bp shorter than Latrodectus α-latrotoxin. All NCBI Accession numbers for these sequences are listed in supplementary table S1, Supplementary Material online.

Maximal amino acid distance among α-latrotoxin sequences was 35.6% (between S. grossa and L. tredecimguttatus), similar to the distance between the functionally distinct species.
paralogs α-latroinsectotoxin and α-latrocrustotoxin from *L. tredecimguttatus* (35.5%) but was no more than 16% different within *Latrodectus* (supplementary table S2, Supplementary Material online). Moreover, 68.7% of all differences between *S. grossa* and *L. tredecimguttatus* α-latrotoxin involved nonconservative changes between different physiochemical classes of amino acids (e.g., between hydrophobic and hydrophilic classes), as determined from Livingstone and Barton (1993) categorization of residue properties (supplementary fig. 1, Supplementary Material online). Within *Latrodectus*, the highest average uncorrected nucleotide distance among α-latrotoxins was 6.1%, which was less than corresponding mt COI distances (13.5%; supplementary table S2, Supplementary Material online). This pattern was reversed when comparing *S. grossa* with *Latrodectus* species, with α-latrotoxin average nucleotide distance (27.9%) being more divergent than average mt COI distance (19.3%; supplementary table S2, Supplementary Material online). Likelihood ratio tests of rate homogeneity were rejected for the 4 kb α-latrotoxin sequences (*P* < 0.01), but could not be rejected for mt COI sequences from the same species (*P*= 0.08). The greater sequence divergence in *S. grossa* versus *Latrodectus* α-latrotoxin relative to mt COI, but the reverse pattern among *Latrodectus*, could be due to α-latrotoxin having an accelerated evolutionary rate outside of *Latrodectus*, and/or a deceleration of its substitution rate within the genus. It should be noted that mt COI shows evidence of saturation in third codon positions outside of *Latrodectus* (Garb et al. 2004), which likely constrains the degree to which it could vary above 19.3% because nearly all changes are restricted to third codon positions. α-Latrotoxin exhibits much greater divergence in first and second codon positions outside of the *Latrodectus* genus, but it is still half as divergent as mt COI within the genus.

α-Latrotoxin domains exhibited increasingly greater divergence from the N-terminal toward the C-terminal domains (wing < body < head < posttranslationally cleaved domains; supplementary table S2, Supplementary Material online). Pairwise nonsynonymous substitutions per nonsynonymous site (*dN*) for α-latrotoxin between *L. tredecimguttatus* versus

**Fig. 1.** (A) Schematic of the primary structure of α-latrotoxin indicating the functional domains, with numbers indicating amino acid position, boxes showing position of 22 ankyrin repeats, and black lines indicating amino (N)- and carboxy (C)-terminal ends that are posttranslationally cleaved to yield the mature protein. (B) (Left) Three-dimensional structure of α-latrotoxin, and (right) homotetramer indicating position of the protein domains determined from cryo-EM microscopy (modified from Ushkaryov et al. 2004). (C) Sliding window analysis of *dN* (nonsynonymous substitutions/nonsynonymous site) between α-latrotoxin from *Latrodectus tredecimguttatus* versus *Steatoda grossa* (black line) and *L. tredecimguttatus* versus *L. geometricus* (dashed line) illustrating variability along gene. Permission for image in fig. 1C received from Elsevier via RightsLink under license number 3078871425016; Publication: **Toxicon**. Title: The multiple actions of black widow spider toxins and their selective use in neurosecretion studies. Type Of Use: reuse in a journal/magazine.
L. geometricus or versus S. grossa, plotted in a sliding window analysis, similarly shows an overall trend of increasing \( d_N \) values going from a 5' to 3' direction, as well as substantial divergence between (in contrast to within) genera (fig. 1C and supplementary table S2, Supplementary Material online). Some pairwise estimates of the ratio of \( d_N \) to synonymous substitutions per synonymous site \( (\omega, d_N/d_S) \) for the \( \alpha \)-latrotoxin head domain among \( \textit{Latrodectus} \) sequences are as much as 1.69, but never exceed 0.508 in any other domain (supplementary table S2, Supplementary Material online). Moreover, no pairwise estimates of \( \omega \) greater than 1 were significantly larger than 1, as determined by codon-based \( Z \)-tests of selection performed in MEGA5 (Tamura et al. 2011), which failed to reject the hypothesis of \( d_N = d_S \) in favor of the alternative hypothesis of \( d_N > d_S \) for any pairwise comparison. All cysteine residues are 100% conserved across the approximately 4.2 kb \( \alpha \)-latrotoxin sequences. In one \( \textit{P. tepidariorum} \) latrotoxin clone, position 393 (a conserved cysteine in all sequences except \( \alpha \)-latrocrustotoxin) was an asparagine; mutation of the cysteine in this position in \( \alpha \)-latrotoxin clone, position 393 (a conserved cysteine in all sequences except \( \alpha \)-latrocrustotoxin) was an asparagine; mutation of the cysteine in this position in \( \alpha \)-latrotoxin largely reflect species relationships previously estimated with all mitochondrial COI codon positions by Garb et al. (2004) as well as with our new data (fig. 3 and supplementary figs. S2 and S3, Supplementary Material online). BLASTx searches against the UniProtKB database found the greatest similarity between \( \alpha \)-latrotoxin and previously reported latrotoxins, followed by ankyrin 1 from \( \textit{Homo sapiens} \) (E score = 2 e\(^{-60}\)) or dTRPA1 (transient receptor potential cation channel subfamily A member 1) from \( \textit{Drosophila} \) when limiting the database to arthropods (supplementary table S3, Supplementary Material online).

Phylogenetic Analyses of \( \alpha \)-Latrotoxin and mt COI

We constructed phylogenetic trees using parsimony, maximum likelihood (ML), and partitioned Bayesian analyses (partitioned by codon position) from nucleotide alignments of the following three data sets: 1) approximately 4 kb of \( \alpha \)-latrotoxin and latrotoxin paralogs from a subset of \( \textit{Latrodectus} \) sequences, and 2) the conserved 618 bp fragment of \( \alpha \)-latrotoxin spanning 28.1% of the wing and 16.4% of the body domains from 41 specimens, and 3) 428–659 bp of mt COI sequences from all specimens, including and excluding third codons positions. The longer latrotoxin sequence alignment (4,233 bp) contained 3,215 variable sites, 2,066 of which were parsimony informative. The Akaike Information Criteria (AIC) in jModeltest selected the GTR + G substitution model for this data set (all model selections for each analysis reported in supplementary table S4, Supplementary Material online). The ML tree (−\( \ln L \)) 26,441.57532; fig. 2), the partitioned Bayesian consensus tree and the single most parsimonious tree (length = 5,741, consistency index [CI] = 0.835, retention index [RI] = 0.734, rescaled consistency index [RC] = 0.612) were identical in topology with high bootstrap values and posterior probabilities (0.93–1.00 for all nodes), respectively. Relationships of \( \alpha \)-latrotoxin largely reflect species relationships previously estimated with all mitochondrial COI codon positions by Garb et al. (2004) as well as with our new data (fig. 3 and supplementary figs. S2 and S3, Supplementary Material online). Within \( \textit{Latrodectus} \), the mactans clade includes all species other than \( \textit{L. geometricus} \) and \( \textit{L. rhodesiensis} \) and these latter two comprise the geometricus clade. Unlike the ML and parsimony analysis, the partitioned Bayesian consensus of all mt COI codon positions did not recover \( \textit{Latrodectus} \) as monophyletic, instead placing the two \( \textit{Steatoda} \) species as sister to the mactans clade to the exclusion of the geometricus clade with a posterior probability of 0.96 (supplementary fig. S2, Supplementary Material online). Monophyly of \( \textit{Latrodectus} \) was also not supported by parsimony, ML and Bayesian analyses of mt COI excluding third codons positions (supplementary figs. S4 and S5, supplementary table S4, Supplementary Material online).
Supplementary Material online), and these trees were less well resolved than mt COI trees including third codon positions. The limited resolution of these trees is likely due to the increase in parsimony-informative characters from 220 in all codons positions to 42 in first and second codons positions alone. Analyses of the 618 bp fragment of $\alpha$-latrotoxin (642 bp in alignment) yielded two ML trees (fig. 4; $-\ln L = 5,853.97981$) that differ only in the arrangement of nearly identical sequences (L. hasselti LhasP4 and LhasP5 relative to the two L. katipo sequences) and both topologies are congruent with the partitioned Bayesian consensus, except for a few nodes with weak support (supplementary fig. S6, Supplementary Material online). Parsimony analyses of this data set resulted in six most parsimonious trees (supplementary fig. S7, Supplementary Material online; length = 5,741, CI = 0.680, RI = 0.723, RC = 0.491). The strict consensus is similar in overall structure to the ML and partitioned Bayesian trees, with sequences falling into well-supported geometricus and mactans clades, but with less resolution among some deeper nodes within the mactans clade. Sequences from P. tepidariorum are highly divergent and not united with the Latrodectus and Steatoda $\alpha$-latrotoxins, indicating they represent $\alpha$-latrotoxin paralogs.

Tests for Selection on $\alpha$-Latrotoxin

The 4 kb and 642 bp $\alpha$-latrotoxin alignments were used to estimate $\omega (d_s/d_o)$ for codons in the alignment and branches of the phylogeny to investigate patterns of selection that have
acted on \( \alpha \)-latrotoxin using three programs: 1) the codeml program in PAML 4.3 (Yang 2007); 2) multiple modules within the HyPhy package via the www.datamonkey.org (last accessed February 1, 2013) web server (Delport et al. 2010); and 3) the package ADAPTSITE (Suzuki et al. 2001). We also examined codon-specific patterns of radical and conservative amino acid replacements in the \( \alpha \)-latrotoxin alignments using ADAPTSITE (Suzuki et al. 2001), as the ratio of these changes can reflect the influence of purifying or positive selection (Suzuki 2007). Collectively, the results of these analyses were consistent in suggesting that \( \alpha \)-latrotoxin has largely been subjected to purifying selection across the majority of its codons and along branches of its phylogeny, with some evidence of positive selection acting on a small subset of codons (tables 1 and 2; fig. 2).

The codeml-free ratio branch model of PAML estimated \( \omega \) along all branches of the 4 kb \( \alpha \)-latrotoxin phylogeny as being far less than 1 (fig. 2). Fixed branch model estimates of \( \omega \) were 0.14 and 0.10 for the 4 kb and 642 bp data sets, respectively, and were a significantly better fit to the data than the branch model fixing \( \omega \) to one, but a significantly worse fit to the free ratio branch model (table 1). The two-ratio branch model for the 4 kb data set, allowing the branch leading to \( \text{Latrodectus} \) \( \alpha \)-latrotoxin to have a different \( \omega \) (0.289) than the background ratio (0.133), was a significantly better fit than the one-ratio model (\( P < 0.005 \)). This suggests an overall history of purifying selection on \( \alpha \)-latrotoxin, with some temporal variation along branches, including elevated \( \omega \) along the \( \text{Latrodectus} \) \( \alpha \)-latrotoxin branch. However, the two-ratio branch model was not significantly different from the one-ratio model for the 642 bp data set. For both the 4 kb and 642 bp alignments, comparisons of the M0 and M3 sites models indicated significant variation in \( \omega \) among codon sites (table 1). The M7 and M8 sites models were significantly

![Fig. 4. ML phylogram from conserved 642 bp alignment (618 bp PCR fragment) of \( \alpha \)-latrotoxin and latrotoxin homologs. Asterisks indicate nodes supported by \( \geq 88\% \) bootstrap support from 100 replicates (ML) or 1,000 replicates (parsimony), and \( \geq 0.95 \) clade posterior support values in partitioned Bayesian analysis. Numbers in parentheses indicate numbers of additional identical sequences (supplementary table S1, Supplementary Material online: 013 L. corallinus identical to 020 L. corallinus and 003 L. diapria; 015 L. mirabilis identical to 014 L. mirabilis and 010 Latrodectus sp.; L. hasselti LhasP4 identical to LhasP5; L. tredecimguttatus NCBI identical to 036 L. tredecimguttatus and Ltre4; L. hesperus LhespUIC identical to Lhesm1 and 049 L. hesperus; 017 L. geometricus identical to 016 L. geometricus; L. geometricus geoCA1 identical to 034 L. geometricus, 033 L. geometricus, 097 L. geometricus and 099 L. geometricus; S. grossa Arizona identical to S. grossa Maui). The tree was rooted with the latrotoxin paralog \( \delta \)-latroinsectotoxin. Partitioned Bayesian and parsimony consensus trees for these data are shown in supplementary figures 6 and 7, Supplementary Material online, respectively.](image-url)
different for the 4 kb alignment (P = 0.02), and although five sites undergoing positive selection were identified with the Bayes empirical Bayes (BEB) procedure, there was insufficient statistical support for this result (all five sites with posterior probability \( \omega > 1 \) being 0.644 or less). Moreover, the M2a model was not a better fit to the data than the M1a model. The MA (modified) branch-sites model for the 4 kb alignment examining evidence for selection at specific sites along the Latrodectus \( \alpha \)-latrotoxin branch was also not significantly different than the null model, fixing this branch to an \( \omega \) of 1. However, the branch-sites model of the 4 kb data set did estimate 11% of sites having a mean \( \omega \) of 2.54 along the Latrodectus \( \alpha \)-latrotoxin branch, and the BEB procedure identified 26 codons as possible targets of positive selection, but only five of the 26 had posterior probabilities between 0.85 and 0.95: 158 L, 400 E, 589 S, 652 F, and 751 W (table 1). The first of these sites is located in the wing domain, while the other four are located in various positions in the body domain. All but the last of these five sites involve changes to amino acids in different physiochemical categories. For the 642 bp alignment, the sites model M8 was not significantly different from M7, nor was the MA (modified) branch-sites model significantly different from the null. However, the BEB procedure following the branch-sites model identified one codon from the body domain (127 T), also identified from the 4 kb branch-sites model (400 E), as having a posterior probability of 0.899 (table 1).

### Table 1. Summary of Codon Substitution Models Examined in Codeml for \( \alpha \)-Latrotoxin, with Estimates of \( \omega \), Model Parameters, Likelihood Values, and Probabilities.a,b

<table>
<thead>
<tr>
<th>Model</th>
<th>In L</th>
<th>Parameter Estimates</th>
<th>( \chi^2 )</th>
<th>Df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Branch 4 kb</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free ratio</td>
<td>-19,296.84</td>
<td>( \omega = 0.143 )</td>
<td>83.26</td>
<td>22</td>
<td>(&lt;0.00001)</td>
</tr>
<tr>
<td>One ratio</td>
<td>-19,338.7</td>
<td>( \omega = 1 )</td>
<td>1788.82</td>
<td>1</td>
<td>(&lt;0.00001)</td>
</tr>
<tr>
<td>Two ratios</td>
<td>-19,334.44</td>
<td>( \omega_0 = 0.133; \omega_1 = 0.289 )</td>
<td>8.07</td>
<td>1</td>
<td>(&lt;0.005)</td>
</tr>
<tr>
<td><strong>Branch 642 bp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One ratio</td>
<td>-3,259.68</td>
<td>( \omega = 0.108 )</td>
<td>417.18</td>
<td>1</td>
<td>(&lt;0.00001)</td>
</tr>
<tr>
<td>Two ratios</td>
<td>-3,259.67</td>
<td>( \omega_0 = 0.108; \omega_1 = 0.100 )</td>
<td>0.01</td>
<td>1</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Sites 4 kb</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>-19,338.47</td>
<td>( \omega = 0.143 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>-19,147.65</td>
<td>( p(0, 1, 2): 0.192, 0.631, 0.177; \omega(0, 1, 2): 0.016, 0.129, 0.501 )</td>
<td>381.63</td>
<td>4</td>
<td>(&lt;0.00001)</td>
</tr>
<tr>
<td>M1a</td>
<td>-19,227.05</td>
<td>( p(0, 1): 0.8568, 0.143; \omega(0, 1): 1.11288, 1.00000 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2a</td>
<td>-19,227.05</td>
<td>( p(0, 1, 2): 0.857, 0.143, 0.000; \omega(0, 1, 2): 0.112, 1.000, 35.731 )</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>M7</td>
<td>-19,153.64</td>
<td>( p = 1.07987; q = 5.23687 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M8</td>
<td>-19,149.98</td>
<td>( p_0 = 0.966; p = 1.241; q = 6.936 (p_1 = 0.035); \omega = 1.000 )</td>
<td>7.32</td>
<td>2</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Sites 642 bp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>-3,259.67</td>
<td>( \omega = 0.108 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>-3,224.98</td>
<td>( p(0, 1, 2): 0.242, 0.617, 0.140; \omega(0, 1, 2): 0.013, 0.117, 0.301 )</td>
<td>69.38</td>
<td>4</td>
<td>(&lt;0.00001)</td>
</tr>
<tr>
<td>M1a</td>
<td>-3,252.75</td>
<td>( p(0, 1): 0.938, 0.062; \omega(0, 1): 0.999, 1.000 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2a</td>
<td>-3,252.75</td>
<td>( p(0, 1, 2): 0.938, 0.062, 0.000; \omega(0, 1, 2): 0.099, 1.000, 21.208 )</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>M7</td>
<td>-3,227.14</td>
<td>( p = 1.25897; q = 9.19405 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M8</td>
<td>-3,227.14</td>
<td>( p_0 = 0.999; p = 1.259; q = 9.194; (p_1 = 0.00001); \omega = 1.000 )</td>
<td>0.00</td>
<td>2</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>Branch-sites 4 kb</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA (mod.)</td>
<td>-19,216.64</td>
<td>( p(0, 1, 2a + 2b): 0.760, 0.120, 0.119; \omega(2a + 2b, foreground): 2.545 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA (fixed)</td>
<td>-19,217.95</td>
<td>( \omega = 1 )</td>
<td>2.61</td>
<td>1</td>
<td>0.106</td>
</tr>
<tr>
<td><strong>Branch-sites 642 bp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA (mod.)</td>
<td>-3,250.71</td>
<td>( p(0, 1, 2a + 2b): 0.922, 0.062, 0.0167; \omega(2a + 2b, foreground): 149.03 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA (fixed)</td>
<td>-3,252.06</td>
<td>( \omega = 1 )</td>
<td>2.68</td>
<td>1</td>
<td>0.100</td>
</tr>
</tbody>
</table>

Note.—BEB positively selected residues (posterior probability \( \omega > 1 \)). M8 4 kb: 468 A 0.644; 524 T 0.596; 711 Q 0.587; 861 N 0.520; 940 S 0.620. Branch-sites, 4.2 kb: 47 A 0.760; 158 L 0.881; 187 E 0.642; 400 E 0.685; 466 T 0.770; 506 S 0.768; 511 F 0.731; 523 V 0.624; 549 Y 0.661; 556 L 0.561; 589 S 0.943; 635 G 0.823; 637 L 0.543; 652 F 0.879; 676 Q 0.554; 697 K 0.580; 712 I 0.568; 739 S 0.548; 751 W 0.873; 857 D 0.547; 892 Q 0.585; 895 R 0.733; 937 S 0.536; 981 I 0.794; 998 H 0.500; 1035 T 0.642. Branch-sites, 642 bp: 127 T 0.899. aPAML run with clean data = 1; \( \omega_1 \) is foreground branch, or branch leading to Latrodectus \( \alpha \)-latrotoxin, whereas \( \omega_0 \) are for background branches. b \( p(0, 1) \) and \( p(0, 1, 2) \) in sites models M3, M1a, and M2a refer to the proportion of sites in each of the two or three estimated \( \omega \) values, respectively; \( p_0 \) and \( p_1 \) in the M8 model refer to the proportion of sites in the two estimated \( \omega \) values; \( p \) and \( q \) in sites models M7 and M8 refer to parameters in the beta distribution \( B(p, q) \) used to estimate \( \omega \) in the interval \([0, 1]\).
for estimating $d_{N}-d_{S}$ at each codon position found no positively selected sites and 130 negatively selected sites (where normalized $d_{N}-d_{S}$ is negative) at a significance level of 0.1. SLAC estimated an overall $\omega$ of 0.28 across all codons. FEL found 10 positively selected codons (two in the wing domain and eight in the body domain; where normalized $d_{N}-d_{S}$ is positive), and 370 negatively selected codons at the 0.1 significance level (table 2 and supplementary fig. S8, Supplementary Material online). FUBAR found no sites under diversifying selection and 639 sites with evidence of negative selection with a posterior probability >0.9. MEME, which allows $\omega$ to vary across codons as well as across branches of the phylogeny, identified 27 codons with evidence of episodic positive selection at a significance level of 0.1. Eight of the 27 codons detected by MEME undergo only nonsynonymous substitutions on the *Latrodectus* $\alpha$-latrotoxin branch (wing domain codons 55 and 158, and body domain codons 313, 468, 660, 725, 754, and 817; table 2). One of these codons (158: from Ala $\rightarrow$ Leu) is among those identified by the BEB procedure in the MA (modified)

<table>
<thead>
<tr>
<th>Codon Positiona</th>
<th>FEL ($d_{N}–d_{S}$)</th>
<th>FEL (P value)</th>
<th>MEME ($\omega^{b}$)</th>
<th>MEME (P value)</th>
<th>Codeml M8 BEB Pr ($\omega &gt; 1$)</th>
<th>Codeml MA(mod) BEB Pr ($\omega &gt; 1$)</th>
<th>ADAPTSITE Radical/Conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-kb Alignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>0.815</td>
<td>0.035</td>
<td>&gt; 100</td>
<td>0.039</td>
<td>—</td>
<td>—</td>
<td>0.424</td>
</tr>
<tr>
<td>63</td>
<td>−0.215</td>
<td>0.584</td>
<td>&gt; 100</td>
<td>0.002</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>102</td>
<td>0.847</td>
<td>0.267</td>
<td>&gt; 100</td>
<td>0.052</td>
<td>—</td>
<td>—</td>
<td>0.244</td>
</tr>
<tr>
<td>158</td>
<td>0.133</td>
<td>0.253</td>
<td>&gt; 100</td>
<td>0.008</td>
<td>—</td>
<td>0.881</td>
<td>0</td>
</tr>
<tr>
<td>173</td>
<td>0.125</td>
<td>0.758</td>
<td>&gt; 100</td>
<td>0.045</td>
<td>—</td>
<td>—</td>
<td>9.312</td>
</tr>
<tr>
<td>181</td>
<td>0.240</td>
<td>0.507</td>
<td>12.337</td>
<td>0.089</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>207</td>
<td>−0.394</td>
<td>0.223</td>
<td>26.499</td>
<td>0.076</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>243</td>
<td>0.347</td>
<td>0.097</td>
<td>&gt; 100</td>
<td>0.104</td>
<td>—</td>
<td>—</td>
<td>1.602</td>
</tr>
<tr>
<td>276</td>
<td>−0.711</td>
<td>(0.051)</td>
<td>&gt; 100</td>
<td>0.045</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>295</td>
<td>−0.288</td>
<td>0.765</td>
<td>24.546</td>
<td>0.023</td>
<td>—</td>
<td>—</td>
<td>0.559</td>
</tr>
<tr>
<td>313</td>
<td>0.186</td>
<td>0.259</td>
<td>&gt; 100</td>
<td>0.094</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>336</td>
<td>0.158</td>
<td>0.213</td>
<td>&gt; 100</td>
<td>0.026</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>468</td>
<td>2.608</td>
<td>0.113</td>
<td>&gt; 100</td>
<td>0.072</td>
<td>0.644</td>
<td>—</td>
<td>0.378</td>
</tr>
<tr>
<td>471</td>
<td>0.289</td>
<td>0.653</td>
<td>12.042</td>
<td>0.039</td>
<td>—</td>
<td>—</td>
<td>0.206</td>
</tr>
<tr>
<td>571</td>
<td>0.613</td>
<td>0.075</td>
<td>&gt; 100</td>
<td>0.109</td>
<td>—</td>
<td>—</td>
<td>3.842</td>
</tr>
<tr>
<td>625</td>
<td>0.359</td>
<td>0.039</td>
<td>&gt; 100</td>
<td>0.010</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>632</td>
<td>0.137</td>
<td>0.685</td>
<td>&gt; 100</td>
<td>0.008</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>655</td>
<td>0.383</td>
<td>0.322</td>
<td>&gt; 100</td>
<td>0.091</td>
<td>—</td>
<td>—</td>
<td>0.358</td>
</tr>
<tr>
<td>657</td>
<td>0.790</td>
<td>0.096</td>
<td>&gt; 100</td>
<td>0.128</td>
<td>—</td>
<td>—</td>
<td>0.54</td>
</tr>
<tr>
<td>660</td>
<td>1.472</td>
<td>0.422</td>
<td>&gt; 100</td>
<td>0.016</td>
<td>—</td>
<td>—</td>
<td>1.876</td>
</tr>
<tr>
<td>683</td>
<td>2.313</td>
<td>0.155</td>
<td>36.496</td>
<td>0.017</td>
<td>—</td>
<td>—</td>
<td>0.468</td>
</tr>
<tr>
<td>691</td>
<td>−0.161</td>
<td>0.897</td>
<td>7.598</td>
<td>0.096</td>
<td>—</td>
<td>—</td>
<td>0.35</td>
</tr>
<tr>
<td>725</td>
<td>0.574</td>
<td>0.093</td>
<td>&gt; 100</td>
<td>0.100</td>
<td>—</td>
<td>—</td>
<td>0.56</td>
</tr>
<tr>
<td>753</td>
<td>−0.123</td>
<td>0.634</td>
<td>&gt; 100</td>
<td>0.065</td>
<td>—</td>
<td>—</td>
<td>2.656</td>
</tr>
<tr>
<td>754</td>
<td>0.476</td>
<td>0.084</td>
<td>&gt; 100</td>
<td>0.100</td>
<td>—</td>
<td>—</td>
<td>0.591</td>
</tr>
<tr>
<td>756</td>
<td>0.444</td>
<td>0.095</td>
<td>&gt; 100</td>
<td>0.111</td>
<td>—</td>
<td>—</td>
<td>1.285</td>
</tr>
<tr>
<td>776</td>
<td>0.907</td>
<td>0.068</td>
<td>&gt; 100</td>
<td>0.090</td>
<td>—</td>
<td>—</td>
<td>0.543</td>
</tr>
<tr>
<td>802</td>
<td>0.090</td>
<td>0.943</td>
<td>&gt; 100</td>
<td>0.023</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>817</td>
<td>0.581</td>
<td>0.080</td>
<td>&gt; 100</td>
<td>0.079</td>
<td>—</td>
<td>—</td>
<td>0.748</td>
</tr>
<tr>
<td>881</td>
<td>0.397</td>
<td>0.254</td>
<td>&gt; 100</td>
<td>0.012</td>
<td>—</td>
<td>—</td>
<td>0.408</td>
</tr>
<tr>
<td>883</td>
<td>0.240</td>
<td>0.559</td>
<td>&gt; 100</td>
<td>0.023</td>
<td>—</td>
<td>—</td>
<td>3.404</td>
</tr>
<tr>
<td>1008</td>
<td>0.213</td>
<td>0.262</td>
<td>&gt; 100</td>
<td>0.092</td>
<td>—</td>
<td>—</td>
<td>0.419</td>
</tr>
<tr>
<td>642 bp Alignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>−0.304</td>
<td>0.524</td>
<td>21.312</td>
<td>0.025</td>
<td>—</td>
<td>—</td>
<td>0.618</td>
</tr>
<tr>
<td>127</td>
<td>0.036</td>
<td>0.886</td>
<td>47.663</td>
<td>0.097</td>
<td>—</td>
<td>0.899</td>
<td>1.336</td>
</tr>
</tbody>
</table>

Note.—Statistical support for positive selection by either method is indicated by underlined P values, value in parentheses indicates site found to be negatively selected with significance for FEL method. Em dash indicates codons not identified as positively selected by the BEB procedure in Codeml.

aSites found to be under positive selection in more than one method are underlined. For 4-kb alignment, codons 55–295 in wing domain, positions 313–883 in body domain, 1,008 in posttranslationally cleaved region; for 642-bp alignment codon 50 in wing domain, codon 127 in body domain.

bValues indicate inferred $\omega (\omega^{b})/\omega^{a}$, values where they exceed 100 include sites where $\omega$ is 0.
branch-sites model of codeml as having $\omega > 1$ on the *Latrodectus* $\alpha$-latrotoxin branch with a posterior probability of 0.88. Another one of the 27 codons (468: from Thr $\rightarrow$ Ala) was identified by the M8 model of codeml as positively selected, without strong statistical support. However, neither of these amino acid substitutions involves substantial changes in charge or hydrophobicity. Of the 32 codons identified in $\alpha$-latrotoxin as positively selected with statistical support using any method, 10 are in the wing domain (comprising 3.3% of wing codons) and 21 are in the body domain (1.1% of all body codons). We also ran the GA-Branch program of the HyPhy package on the 4 kb alignment, which partitions branches in a phylogeny into multiple $d_{\text{NS}}/d_{\text{S}}$ rate classes. The GA-Branch analysis on the 4 kb $\alpha$-latrotoxin tree yielded three $d_{\text{NS}}/d_{\text{S}}$ rate classes: 0.146 over 76% of the tree; 0.232 over 23% of the tree; and 0.475 over 1% of the tree. The *Latrodectus* $\alpha$-latrotoxin branch falls into this middle $d_{\text{NS}}/d_{\text{S}}$ rate class of 0.232, which is nearly identical to the free ratio branch model estimate of 0.238 determined for this branch by codeml. The SLAC, FEL, FUBAR, and MEME methods were also applied to the 642 bp $\alpha$-latrotoxin alignment, and only two codons (codon 50 in the wing and 127 in the body domain) were detected by MEME as having experienced positive selection (table 2). SLAC, FEL, and FUBAR identified 36, 59, and 67 total codons, respectively, as negatively selected. Codon 50 was also identified by the MEME procedure as positively selected in the 4 kb alignment (codon 295), whereas codon 127, though not identified in the 4 kb HyPhy analyses, was identified by the BEB procedure in codeml following the branch-sites model for both the 642 bp and 4 kb alignments (codon 400; table 2).

We calculated the rates of various types of nonsynonymous substitutions with the adaptite-p and adaptite-t programs in the ADAPTSITE package to estimate radical substitutions (cr) per radical site (sr), as well as conservative substitutions (cc) per conservative site (sc) at each codon in both alignments (Suzuki et al. 2001). Radical and conservative nonsynonymous substitutions are divided by the difference in charge of the amino acids changed (Suzuki 2007). For the 4 kb alignment, the adaptite-p analysis showed that the radical over conservative substitution rate ratio [(cr/sr)/(cc/sc)] exceeded 1 at 231 (21.9%) of the 1,051 total codons in the $\alpha$-latrotoxin alignment (an additional 114 codons with radical substitutions had 0 conservative substitutions). The ratio was less than 1 in 561 (53.3%) of total codons, indicating a greater number of conservative substitutions. Tests of neutrality by adaptite-t found 92 sites were under negative selection based on the $d_{\text{NS}}/d_{\text{S}}$ ratio, and five codons were under negative selection using the radical/conservative substitution ratio, but no codons were identified as positively selected with statistical support using either criterion. Similarly, analyses of the 642 bp alignment found that (cr/sr)/(cc/sc) exceeded 1 at 28 (22.0%) of the 127 examined codons, an additional 12 codons with radical substitutions had no conservative substitutions, and (cr/sr)/(cc/sc) was less than 1 in 71 (55.9%) of the codons. Twenty-four codons were negatively selected based on the $d_{\text{NS}}/d_{\text{S}}$ ratio, and none were positively selected either by the $d_{\text{NS}}/d_{\text{S}}$ or (cr/sr)/(cc/sc) ratios.

The program TreeSAAP v.3.2 (Woolley et al. 2003; McClellan and Ellison 2010) was also used to detect positive selection on the 4 kb $\alpha$-latrotoxin alignment based on physiochemical changes in protein sequences. The sliding window analysis in TreeSAAP identified 15 physiochemical categories in which $\alpha$-latrotoxin had a significant Z score ($\geq 3.09$) for a magnitude 8 (the most extreme) change, indicating evidence of positive selection. Only the following three categories were significant for larger windows ($\geq 20$ bp windows): isoelectric point, equilibrium constant, and mean root mean square (r.m.s.) fluctuation displacement. Significance scores for these three categories, along with surrounding hydrophobicity, were plotted along the length of $\alpha$-latrotoxin (supplementary fig. S9, Supplementary Material online), which indicate significant changes in hydrophobicity concentrated in the $\alpha$-latrotoxin wing domain, whereas significant changes in isoelectric point and mean r.m.s. fluctuation displacement are concentrated in the body and head domains. Sites affected by positive selection, when mapped on the *Latrodectus* branch in the $\alpha$-latrotoxin phylogeny, affected 4, 10, or 1 residues in the wing, body, and head domains, respectively (supplementary table S5, Supplementary Material online). Two of these residues (554 and 760) were identified by the codeml M8 model as positively selected, but without statistical support (BEB posterior probability $\omega > 1$ between 0.587 and 0.596).

**Discussion**

$\alpha$-Latrotoxin Distribution, Diversity, and Biomedical Significance

$\alpha$-Latrotoxin is a fundamental tool in the study of vertebrate neurosecretion and is the molecule responsible for the suite of symptoms (latrotoxicism) resulting from black widow spider envenomation in humans. Despite its biomedical importance, research on this toxin has overwhelmingly focused on the functional mechanisms of $\alpha$-latrotoxin from *L. tredecimguttatus*. Recently, Graudins et al. (2012) published a second $\alpha$-latrotoxin sequence from *L. hasselti* (red-back spider), and provided evidence of its presence (along with other latrotoxins) from short peptide sequences obtained through mass spectrometry from two other *Latrodectus* species (*L. mactans* and *L. hesperus*). The limited sequence data have precluded molecular evolutionary analyses of $\alpha$-latrotoxin, or a determination of its wider phylogenetic distribution. In this study, we provide evidence of $\alpha$-latrotoxin from a larger sampling of *Latrodectus* species, including members of the divergent geometricus clade (containing brown widows), as well as from species in the genus *Steatoda*. We have also isolated fragments of latrotoxin paralogs from the theridiid species *Parasteatoda tepidariorum* (common house spider). Morphological- and molecular-based phylogenies of the family Theridiidae show *Steatoda* as a possible sister genus of *Latrodectus*, whereas *Parasteatoda* is in a distantly related subfamily within Theridiidae (Aagnarsson 2004; Arnedo et al. 2004). This distribution indicates the widespread

Evolution of Vertebrate Toxin in Black Widow Venom · doi:10.1093/molbev/mst011 MBE

1007
occurrence of latrotoxins across this extremely large spider family (2,351 spp; Platnick 2013), and places the origin of  ɛ-latrotoxin at the common ancestor of Latrodectus and Steatoda or earlier.

Although the cDNA data presented here show the expression of ɛ-latrotoxin in Steatoda venom, the latrotoxin paralogs identified from Parasteatoda were isolated from genomic DNA, and it is unknown whether they encode venom components, or proteins performing different functions in other tissues. Some clinical and neurophysiological work suggests the presence of latrotoxins in Parasteatoda venom (Gillingwater et al. 1999; Isbister and Gray 2003a). Most theridiid species are much smaller than Latrodectus species (Ubick et al. 2005), and are rarely the cause of human envenomation. Further investigation into the venoms of these numerous, more cryptic theridiid species will provide more detailed insights into the origin and functional diversification of latrotoxins.

The lack of stop codons, limited variability, conserved cysteine residues, similar lengths, and posttranslational processing signals of Latrodectus ɛ-latrotoxin sequences, suggest that all species in the genus express a form of this vertebrate-specific toxin with strong functional similarity to the ortholog from L. tredecimguttatus. This finding has important implications for treating widow spider bites globally (Graudins et al. 2002, 2012; Daly et al. 2007). Black widow spiders are one of the two most clinically significant types of spiders worldwide (Vetter and Isbister 2008), with at least 5,000 envenomations reported annually in Australia for the red-back spider alone (Isbister and White 2004). The limited sequence variability of ɛ-latrotoxin across the mactans clade of Latrodectus (maximum pairwise nucleotide distance = 5.8%), explains why redback spider antivenom (RBSAV) is broadly effective in treating bites from other Latrodectus species (Daly et al. 2007; Graudins et al. 2012).

In contrast to the high conservation of ɛ-latrotoxin across black widow species, the level of protein distance between S. grossa and all other Latrodectus ɛ-latrotoxins (as much as 35.6%) equals the distance between functionally distinct latrotoxin paralogs characterized from L. tredecimguttatus (34.7% between ɛ-latrotoxinsectotoxin and ɛ-latrocristotoxin), and 68.7% of these differences involve nonconservative amino acid substitutions. This greater sequence divergence, as well the substantial divergence in ɛ-latrotoxin between the mactans clade and the geometricus clade, most likely explains why bites from both Steatoda species and L. geometricus are less severe than bites from species of the mactans clade (Muller 1992; Graudins et al. 2002). Specifically, L. geometricus bites elicit neurologically similar symptoms (pain and nausea) to bites from Latrodectus species in the mactans clade, but L. geometricus bites are usually not as debilitating (Muller 1992). Bites from some Steatoda species can also produce local pain and nausea, but are far less severe than bites from Latrodectus species, including L. geometricus (Graudins et al. 2002). However, RBSAV has also been effective in neutralizing the effect of S. grossa venom in humans, and experimentally in vertebrate animal preparations (Graudins et al. 2002, 2012). Because ɛ-latrotoxin is the only component from black widow spider venom associated with neurotoxic effects in vertebrates, this suggests RBSAV is binding to some epitopes in S. grossa ɛ-latrotoxin despite its substantial divergence from Latrodectus ɛ-latrotoxin.

α-Latrotoxin Phylogeny and Evolutionary Origins

The phylogeny of ɛ-latrotoxin provides a framework for examining changing patterns of selection in this venom toxin among lineages over time, and also contributes to our understanding of relationships among Latrodectus species. Several species in the Latrodectus mactans clade are difficult to diagnose using morphological characters, and for many years, multiple geographically widespread species were synonymized as L. mactans (Levi 1959). Recent analyses of mitochondrial DNA identified substantial genetic divergence among these previously synonymized species and determined various levels of species relationships (Garb et al. 2004; Griffiths et al. 2005; Vink et al. 2008). More molecular markers are needed to fully resolve the Latrodectus phylogeny and rapidly evolving markers will be especially important for tracking the spread of the invasive species L. geometricus and L. hasselti (Brown et al. 2008; Vincent et al. 2008; Vink et al. 2008). ɛ-Latrotoxin’s lack of introns and its limited variability within species (e.g., identical sequences between North American and Argentine L. geometricus) suggests it may not be especially helpful for determining population structure without a more thorough sampling of allelic diversity. However, it does hold promise as a marker for diagnosing species and determining their relationships. The ɛ-latrotoxin gene tree provides additional support for some nodes previously identified with mitochondrial COI (Garb et al. 2004), such as the split between the geometricus and mactans clades, as well as much stronger support for monophyly of Latrodectus than does mt COI. However, relationships among major lineages within the mactans clade exhibit discordance between the mt COI and ɛ-latrotoxin trees (figs. 2–4). For example, mt COI shows L. tredecimguttatus as an early branching lineage, and L. pallidus as more closely related to the South American Latrodectus, whereas ɛ-latrotoxin shows North American Latrodectus species as an early branching lineage, but L. tredecimguttatus being more closely related to the South American species (figs. 3 and 4). In any cases of discordance, those branches did not receive significant support, and the discordance is likely due to insufficient phylogenetic signal in one or both of these genes. Moreover, caution should be exercised when using members of gene families encoding venom toxins for generating species trees, as it may be difficult to identify orthologs and detect gene loss (Casewell et al. 2011). However, as of yet, there is no evidence for highly similar copies of ɛ-latrotoxin within Latrodectus genomes.

ɛ-Latrotoxin is most likely descended from a gene duplicate of an ancestral latrotoxin with insecticidal function, but determining the closest relative of ɛ-latrotoxin requires additional homologs to confidently root the gene family tree. BLAST searches of L. tredecimguttatus ɛ-latrotoxin against NCBI and UniProtKB databases did not uncover sequences similar enough to latrotoxins to be incorporated in the
phylogenetic analyses. Holz and Habener (1998) identified two short (10–12 amino acid) regions of $\alpha$-latrotoxin similar to extendin-4, a toxin isolated from the venom of the Gila monster Heloderma suspectum. Extendin-4 is in the family of secretagogic hormones that includes glucagon-like peptide 1 (GLP-1). The extracellular domain of GLP-1’s receptor (GLP1-R), which extendin-4 is an agonist of, also shares sequence similarity to a short region of latrophilin (CIRL), one of the receptors of $\alpha$-latrotoxin. However, the similarity of $\alpha$-latrotoxin to extendin-4 is considered the result of molecular mimicry (convergence) of the authentic ligand of latrophilin and GLP1-R, respectively (Holz and Habener 1998).

BLAST searches conducted for this study found greatest similarity of $\alpha$-latrotoxin to different ankyrin motif (ANK) rich proteins, including ankyrin 1 and ankyrin 2,3/unc 44 as well as to Drosophila transient receptor potential cation channel subfamily A member 1 (dTTRPA1) (supplementary table S3, Supplementary Material online). Both ankyrin 1 and ankyrin 2,3/unc 44 are members of a protein family containing 14–20 ANK repeats that link integral cell membrane proteins to the spectrin-based membrane skeleton (Bennet and Baines 2001), but are substantially longer than latrotoxins. When limiting BLAST searches to arthropod sequences, $\alpha$-latrotoxin shows greatest similarity to dTRPA1, a sequence of similar length to $\alpha$-latrotoxin with 16–17 ANK repeats. dTRPA1, also known as the Wasabi receptor, functions as a calcium permeable transmembrane channel primarily expressed in sensory neurons and underlies sensitivity to temperature and chemical irritants (Cordero-Morales et al. 2011). Given $\alpha$-latrotoxin’s activity in forming a calcium permeable ion channel, this result suggests another potential origin of latrotoxins as paralogs of endogenous spider TRPA1, which is yet to be characterized.

Ecological Implications of $\alpha$-Latrotoxin Molecular Evolution

Venom molecules are widely known for their rapid evolution (Duda and Palumbi 1999; Gibbs and Rossiter 2008), and the unexpected conservatism of $\alpha$-latrotoxin across black widow species contrasts with its divergence from the brown widow ortholog, and much greater divergence from the Steatoda ortholog. This divergence in $\alpha$-latrotoxin’s primary sequence likely explains the functional variation observed across these species’ venoms. In comparison with Latrodectus species from the mactans clade, bites from L. geometricus and S. grossa are less severe and their venoms have greater LD$_{50}$ (median lethal dose) values in mice, indicating a reduced vertebrate toxicity relative to black widow venom (Muller et al. 1989, 1992). Moreover, venoms from L. geometricus and Steatoda are unable to elicit as much neurotransmitter release from vertebrate neurons as do venoms from species in the mactans clade, which contains black widows and red-back spiders (Muller et al. 1989, 1992; Graudins et al. 2002, 2012). Yet, bites from L. geometricus still produce more severe symptoms than do bites from Steatoda species, and L. geometricus venom elicits more neurotransmitter release from vertebrate neurons than does venom from Steatoda species. The phylogenetic relationships among these species suggest a shift toward increased vertebrate toxicity in the venom of Latrodectus, and particularly in the mactans clade. However, establishing the reduced vertebrate toxicity, and possible greater insecticidal activity of S. grossa $\alpha$-latrotoxin will require functional assays of recombinant toxin.

Across Latrodectus and Steatoda, the dominant mode of selection operating on $\alpha$-latrotoxin appears to be purifying selection, as estimates of $\omega$ using multiple methods found the majority of codons and branches of the phylogeny with $\omega$ < 1. However, several methods (FEL, MEME, codeml M8 sites, and MA [modified] branch-sites models) identified a small fraction of codons as having been subjected to positive selection, and TreeSAAP analyses also detected evidence of positive selection on $\alpha$-latrotoxin for several physiochemical/biochemical characteristics. In addition, both the codeml MA (modified) branch-sites model and the GA-Branch method of HyPhy suggested a significant increase in $\omega$ along the Latrodectus $\alpha$-latrotoxin branch and some pairwise estimates of $\omega$ for $\alpha$-latrotoxin’s head domain exceeded 1, suggesting past episodes of more rapid evolution, as well as variation among domains and codons in their functional constraints. These patterns suggest an earlier, more rapid evolution of $\alpha$-latrotoxin, followed by stronger purifying selection, coincident with the origin of the Latrodectus genus and a probable shift in $\alpha$-latrotoxin toward increased vertebrate toxicity.

A shift toward increased vertebrate toxicity in black widow spider venom may be explained by changes in the ecology of these species. Female theridiids range in total body size from 0.8 to 13.0 mm, with Latrodectus females being the largest in size for the family (6–13 mm; Kaston 1970; Ubick et al. 2005). Latrodectus females also build large, robust webs that can intercept and retain small vertebrates. Although widow spiders are generalist predators of arthropods, there are many published accounts of these spiders consuming small vertebrate prey such as geckos, small lizards, snakes, and mice (McCormick and Polis 1982; Hodar and Sanchez-Pinero 2002). Immobilization of vertebrate prey by widow spiders is undoubtedly facilitated by $\alpha$-latrotoxin’s presence in their venom. The possibility that $\alpha$-latrotoxin could serve a defensive function for black widows is suggested by the apparent aposematic coloration of the females, which in the mactans clade have varying levels of bright red pigmentation on top of a contrasting black background on their abdomens (Kaston 1970). Brown widow females and species in closely related genera have much less contrast in their color patterning or are a relatively uniform color (fig. 3).

Long-term purifying selection acting on Latrodectus $\alpha$-latrotoxin implies that it is a functional venom component, and that the maintenance of this vertebrate neurotoxin must have important fitness consequences. It could be that $\alpha$-latrotoxin is utilized by widow spiders in vertebrate predation or in defense from vertebrates more frequently than previously recognized. Another possibility is that in addition to vertebrate toxicity, Latrodectus $\alpha$-latrotoxin also has an insecticidal function. $\alpha$-latrotoxin’s lack of insecticidal activity was determined from a small number of species (locust, blowfly, or cabbage looper moth larvae; Knipper et al. 1986; Magazanik et al. 1992; Kiyatkin et al. 1995). It is possible
that neurophysiological application of $\alpha$-latrotoxin to other insects might indicate a broader phyletic spectrum of toxicity. Finally, $\alpha$-latrotoxin may interact with other venom components in such a way that increases the insecticidal activity of the venom cocktail. Although this study reveals a history of persistent selection on $\alpha$-latrotoxin in black widow spider venom, further investigation of its function in nature is needed to understand its role in the ecology of widow species.

The hypothesis that $\alpha$-latrotoxin was driven by positive selection during the early evolution of Latrodectus, but that advantageous alleles useful in prey acquisition or defense were subsequently maintained by purifying selection, is one of several scenarios that could explain the observed patterns of variability in this gene. For example, relaxation of purifying selection, as opposed to positive selection, due to a population bottleneck during the origin of Latrodectus, is an alternative possibility to explain the increased rate of non-synonymous substitutions on the $\alpha$-latrotoxin Latrodectus branch. Background selection, where purifying selection is acting on a gene linked to $\alpha$-latrotoxin, could also account for the limited variation in $\alpha$-latrotoxin across black widow species. Given the lack of similar studies of Latrodectus nuclear genes, it is difficult to comprehensively evaluate these hypotheses. It will therefore be important to consider patterns of selection on other black widow spider genomic regions, to further determine whether molecular evolution of the vertebrate neurotoxin $\alpha$-latrotoxin is atypical relative to other venom and non-venom genes. More population-level molecular and ecological studies of Latrodectus species are also needed to illuminate their demographic histories, such as historical changes in population size.

This article provides a critical step in identifying the phylogenetic range of species harboring the dangerous human neurotoxin $\alpha$-latrotoxin. Moreover, we provide a substantially greater understanding of the variability and evolution of $\alpha$-latrotoxin, which is essential for predicting the spectrum of spider bites that can be effectively treated with current antivenoms and broadens the available set of probes for studies of vertebrate neurosecretion.

Materials and Methods

Taxon Sampling

A total of 38 individuals representing 16 Latrodectus species were sampled to cover broad phylogenetic and geographic diversity across the genus. We also sampled two species from the putative sister genus Steatoda, as well as a more distantly related theridiid, Parasteatoda tepidariorum (common house spider) (Agnarsson 2004; Arnedo et al. 2004). A complete list of specimens, species identification and collecting localities are in supplementary table S1, Supplementary Material online.

Genomic PCR

Two sets of $\alpha$-latrotoxin sequences were amplified from genomic DNA from the sampled specimens. Genomic DNA was extracted from each specimen using the QIAGEN DNeasy Tissue kit (QIAGEN, Inc.). For the first set, nearly the entire $\alpha$-latrotoxin coding sequence (~4 kb) was amplified from eight specimens representing divergent Latrodectus species. Sequences were amplified in successively overlapping fragments, using a total of 2–23 primer pair combinations per species. All primer sequences, primer combinations used, and PCR conditions used in this study are listed in supplementary tables S6 and S7, Supplementary Material online. Reactions were primarily conducted using nested PCR reactions, with an initial touchdown PCR reaction, followed by amplification of the product using internal nested primers. A second set of sequences spanning 618 bp of $\alpha$-latrotoxin’s wing and body domains were collected from 41 specimens (from 18 Latrodectus and two Steatoda species). Because the directly sequenced PCR product from Parasteatoda tepidariorum yielded a poor-quality chromatogram with messy peaks and many ambiguous bases, it was cloned into pCR2.1-TOPO plasmids (Invitrogen). Plasmid inserts were sequenced and two $\alpha$-latrotoxin paralogs were recovered. A 428–659 bp fragment of the mt COI was amplified and sequenced from all specimens using primers and conditions reported in Garb et al. (2004). All latrotoxin and COI PCR products were directly sequenced in both directions.

RT-PCR and Inverse PCR

Genomic PCR amplification of $\alpha$-latrotoxin using the designed primers was largely unsuccessful outside of Latrodectus. To complete the $\alpha$-latrotoxin coding sequence from Steatoda grossa, a portion of the conserved region that could be amplified with genomic PCR was used to design primers for RT-PCR of the downstream region using methods modified from Binford et al. (2009). First, venom was collected from 12 S. grossa individuals using electro-stimulation to promote gene transcription in venom glands. Venom glands were then dissected from these individuals and flash frozen at −80°C two days after venom collection. Total RNA was extracted from the glands by homogenization with TRIzol (Invitrogen). Single-stranded (ss) cDNA was synthesized from venom gland RNA, using the SuperScript III reverse transcriptase protocol (Invitrogen), with an adaptor oligo-dT primer targeting the 3'-end of mRNAs (3' RACE Adapter 5'-ggccacgcgtcgactagactttttttttttttttttt-3'). This ss cDNA (1:10 dilution) was used as a PCR template with gene-specific forward primer 1,781-f (5'-ccctaaagccctgcttgatagga-3'), a primer matching cDNA 3' ends (3'UniPrime 5'-ggccacgcgtcgactagactag-3'), and AccuPrime Taq DNA Polymerase (Invitrogen). Resulting bands were gel-purified and cloned using the pCR4-TOPO vector (Invitrogen). Positive clones were PCR screened, purified and four were sequenced in their entirety.

To obtain the 5' end of $\alpha$-latrotoxin from S. grossa, inverse (i) PCR was performed with high molecular weight DNA from one S. grossa using the protocol of Sambrook and Russell (2000). Aliquots of the DNA were separately digested with the restriction enzymes, XbaI, PstI, and HindIII. Each completed digest was incubated with T4 DNA ligase to promote circularization. The self-ligated DNAs were templates in nested PCR reactions with primer pairs directed away from each other to amplify the unknown
encoded by the NCBI protein database and Swissprot, or with tBLASTx BLAST searches (BLASTx and BLASTp with psiBLAST) against the NCBI EST database (dbEST). Searches were performed with entire databases and also limiting the database against specific sites along the α-latrotoxin branch connecting to crown Latrodectus to have a variable foreground ratio (Yang 1998). The MA (modified) branch-sites Model A was also compared with the MA (fixed) model to detect positive selection that may have affected specific sites along the α-latrotoxin branch connecting to crown Latrodectus (Yang and Nielsen 2002). Statistical differences among models were estimated using a likelihood ratio test, comparing twice the difference between log-likelihood scores under each model.

In the HyPhy package via datamonkey.org, we ran the SLAC, FEL and FUBAR modules that use differing methods to estimate \( d_{sl}/d_s \) at every codon in both the 4 kb and 642 bp α-latrotoxin alignments and report which codons show evidence of positive or negative selection, using default significance levels. We also ran the MEME module to detect diversifying episodic selection acting on codons, and the GA-Branch module to fit \( d_{sl}/d_s \) rate classes to branches of the tree (Delport et al. 2010). For all modules of HyPhy, analyses were run using the nucleotide evolution model selected by the AIC in the datamonkey model selection tool, and providing the ML tree topology derived from the corresponding alignment.

We ran the adaptsite-p and adaptsite-t programs in the ADAPTSITE package to estimate the \( d_{sl}/d_s \) rate ratio and the radical/conservative \([(cr)/ss]/(cc/sc)] \) amino acid substitution rate ratio at each codon in both the 4 kb and 642 bp α-latrotoxin alignments and to determine the probability of those values under selective neutrality (Suzuki et al. 2001). Following the program requirements, all codons in the alignment with gaps and ambiguities were removed (equivalent to cleanData = 1 matrix used in codeml analyses) and a neighbor-joining tree, constructed by the program njtree was supplied. Substitutions were estimated based on the transition/transversion rate ratio of 2.2 estimated from the α-latrotoxin alignment in codeml of PAML.

The 4 kb α-latrotoxin alignment was also analyzed using the program TreeSAAP v.3.2 (Woolley et al. 2003; McClellan and Ellison 2010), but excluding α-latrotoxin paralogs. A sliding windows analysis of eight magnitude categories for changes in physicochemical properties was conducted using a window size of 20 sites wide taken over each site. All magnitude category 6–8 changes with \( P \) values \( \leq 0.001 \), indicating a significant physicochemical change and positive selection, were mapped over the length of the sequence and each significant site was mapped onto the Latrodectus α-latrotoxin branch.

To identify possible latrotoxin homologs in other organisms, α-latrotoxin sequences were also subjected to various BLAST searches (BLASTx and BLASTp with psibLAST) against the NCBI protein database and Swissprot, or with tBLASTx against the NCBI EST database (dbEST). Searches were performed with entire databases and also limiting the database to Arthropoda.
Supplementary Material

Supplementary tables S1–S7 and figures S1–S9 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank the following individuals for providing the specimens used in this study: M. Andrade, S. Crews, R. Gillespie, A. Gonzalez, M. Harvey, M. Hedin, P. Krushelnycky, C. Kristensen, A. Leroy, L. Lotz, Y. Lubin, M. Segoli, R. Vetter, and C. Vink. The authors also thank N. Ayoub, G. Binford, and K. Zinsmaier for providing advice and assistance in data collection. A. Lancaster provided assistance with the sliding window analysis and ADAPTSITE. They also thank N. Ayoub, M. McGowen, and J. Starrett for critical review of this manuscript. This work was supported by the National Institutes of Health grants (NIGMS 1F32GM083661-01 and 1R15GM097714-01) to J.E.G.

References


San Diego Soc Nat Hist. 1633–82.


